Physical & Occupational Therapy In Pediatrics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ipop20

Measuring Early Problem-Solving in Young Children with Motor Delays: A Validation Study

Rebecca M. Molinini , Natalie A. Koziol , Tanya Tripathi , Regina T. Harbourne , Sarah Westcott McCoy , Michele A. Lobo , James Bovaird & Stacey C. Dusing

To cite this article: Rebecca M. Molinini , Natalie A. Koziol , Tanya Tripathi , Regina T. Harbourne , Sarah Westcott McCoy , Michele A. Lobo , James Bovaird & Stacey C. Dusing (2021): Measuring Early Problem-Solving in Young Children with Motor Delays: A Validation Study , Physical & Occupational Therapy In Pediatrics, DOI: 10.1080/01942638.2020.1865501

To link to this article: https://doi.org/10.1080/01942638.2020.1865501

Published online: 01 Feb 2021.

PHYSICAL & OCCUPATIONAL THERAPY IN PEDIATRICS
https://doi.org/10.1080/01942638.2020.1865501

Measuring Early Problem-Solving in Young Children with Motor Delays: A Validation Study

Rebecca M. Molinini[a], Natalie A. Koziol[b], Tanya Tripathi[c], Regina T. Harbourne[d], Sarah Westcott McCoy[e], Michele A. Lobo[f], James Bovaird[g], and Stacey C. Dusing[h*]

[a]Motor Development Lab, Virginia Commonwealth University, Richmond, Virginia, USA; [b]Nebraska Center for Research on Children, Youth, Families and Schools, University of Nebraska-Lincoln, Lincoln, Nebraska, USA; [c]Pediatric and Rehabilitation Laboratory, The Ohio State University, Columbus, Ohio, USA; [d]Rangos School of Health Sciences, Duquesne University, Pittsburgh, Pennsylvania, USA; [e]Department of Rehabilitation Medicine, University of Washington, Seattle, Washington, USA; [f]Move to Learn Innovation Lab, Department of Physical Therapy and Biomechanics and Movement Science Program, University of Delaware, Newark, Delaware, USA; [g]Educational Psychology, University of Nebraska-Lincoln, Lincoln, Nebraska, USA; [h]Department of Physical Therapy, Virginia Commonwealth University, Richmond, Virginia, USA

ABSTRACT

Aim: There is a lack of evidence-based tools for measuring problem-solving in young children with motor delays. The purpose of this study was to evaluate the construct validity and responsiveness of the Assessment of Problem-Solving in Play.

Methods: 125 young children (10.72, SD 2.62 months) with mild, moderate, and severe motor delays were assessed with the Bayley Scales of Infant and Toddler Development, Third Edition Cognitive Scale and the Assessment of Problem-Solving in Play up to 4 times over 12 months. The baseline and change over time assessment scores were compared.

Results: The Assessment of Problem-Solving in Play was strongly, positively correlated with the Bayley Scales of Infant and Toddler Development, Third Edition Cognitive Scale raw scores at baseline (r=.83, p<.001) and for changes in scores across time (r=.64, p<.001). On average, participants demonstrated positive change in problem-solving scores across time. Participants with severe motor delay scored lower at baseline and changed less as compared to other participants.

Conclusions: Results provide evidence for the construct validity and responsiveness of the Assessment of Problem-Solving in Play scores in quantifying problem-solving in young children with motor delays 7-27 months of age.

ARTICLE HISTORY
Received 18 February 2020
Accepted 14 December 2020

KEYWORDS
Problem solving; infant development; early intervention; psychometric testing; motor development; perceptual-cognitive development; Developmental delay; cerebral palsy

Problem-solving is an important outcome in young children related to both cognitive and social-emotional development (Gibbs & Teti, 1990; Walker et al., 2008). Early problem-solving is manifested as the motor expression of the integration of a number of fundamental cognitive skills, such as attention, memory, perception, and action (Greenwood et al., 2006; Walker et al., 2008). The development of problem-solving is

CONTACT Stacey C. Dusing stacey.dusing@pt.usc.edu Division of Biokinesiology and Physical Therapy, University of Southern California, Los Angeles, CA 90033, USA.

*Dr. Dusing's affiliation changed to the University of Southern California during the review process. However, her role in data collection and funding were completed at Virginia Commonwealth University.

heavily grounded in early motor experiences as motor exploration provides the means through which children encounter new problems, practice and refine problem-solving strategies (Harbourne & Berger, 2019; Walker et al., 2008). Early motor exploration, such as mouthing and banging, allows infants to learn about body and object affordances. Over time infants refine their motor exploration, leading to more goal-directed and sophisticated play with toys. Children with motor delays are afforded fewer self-initiated opportunities to explore and learn from their environment, putting them at risk for delays in problem-solving and other domains of cognition (Adolph & Franchak, 2017; Gibson, 1988). Early identification of delays in problem-solving is pertinent to initiating early intervention services, monitoring children's response to intervention, and determining if the child is on a developmental trajectory toward school readiness (Greenwood et al., 2006; Walker et al., 2008). Measuring problem-solving in young children is essential as these skills serve as the foundation for later success in academics and interpersonal relationships (Gibbs & Teti, 1990; Lobo et al., 2013; Walker et al., 2008).

Standardized developmental assessments are a commonly used metric of cognition in children. However, these assessments typically require that they be completed in a controlled environment by a trained examiner using a standardized set of tools and instructions (Lobo & Galloway, 2013; Yue et al., 2019). Most cognitive developmental assessments in children measure cognition as a whole, and do not provide information on individual cognitive domains, such as problem-solving (Morgan et al., 2019). In addition, many of the cognitive items used for assessment of young children are discrete tasks that are dependent on motor skills to execute. For example, the speed in which a child puts pegs in a board, completes a puzzle, or finds a hidden object all contribute to their scores on multiple cognitive assessments (Barnett et al., 2004; Bayley, 2006; Newborg et al., 1984). The impact of motor delay or limited motor skills reduces the validity of many standardized measures of early learning (Lobo & Galloway, 2013; Morgan et al., 2019; O'Grady & Dusing, 2015).

Although standardized assessment tools often have good psychometric properties, they may not capture the abilities of children in real world situations (Lobo & Galloway, 2013; O'Grady & Dusing, 2015). In contrast, play-based assessments quantify child-directed activities and provide a representation of the way children use their skills in their day-to-day activities (Lobo & Galloway, 2013; O'Grady & Dusing, 2015). Play-based assessments can be administered quickly, frequently, and in coordination with traditional standardized measures (Lobo & Galloway, 2013; O'Grady & Dusing, 2015; Walker & Greenwood, 2010). In response to criticisms of standardized assessments and the mandate for early intervention programs to monitor progress in multiple domains, a number of progress-monitoring play-based measurement systems have recently been developed. Examples include the Infant and Toddler Individual Growth and Development Indicators (IGDIs), myIGDIs for preschool, and Dynamic Indicators of Basic Early Literacy Skills for K-3 (Greenwood et al., 2011).

The purpose of this paper is to highlight the development and validation of a tool that tracks problem-solving skills over time in children with motor delays. First, a problem-solving tool designed to monitor young children's development will be described. Second, the early use of this tool and the subsequent modification process will be

reported. Third, the study on the validation of a new measure for tracking problem-solving in young children with motor delays will be presented.

Early Problem-Solving Indicator

The Early Problem-Solving Indicator (EPSI) is one component of the Infant and Toddler IDGIs (Greenwood et al., 2006; Juniper Gardens Children's Project, 2017) and quantifies problem-solving, by scoring live or from a single viewing of a video, the frequency in which children perform problem-solving skills of varying difficulties during a naturalistic play assessment. The problem-solving skills, from easiest to most difficult, include: Look, Explore, Function, and Solution. The EPSI was developed in 2006 and tested on 30 children with typical development aged 12-48 months and has moderate criterion validity ($r = .42$) with the Bayley Scales of Infant Development-II Mental Development Index (Greenwood et al., 2006).

Modification Process

Following training on the EPSI and discussion with the developers, our team began using the EPSI to evaluate early problem-solving in children with or at risk for motor delay. Our target population was different from the developers in that they were younger (as young as 3 months corrected age) and all participants had a known motor delay or were at high risk for motor delay due to extreme prematurity (Dusing et al., 2015, 2019; Harbourne & Berger, 2019). Due to the novel use of the EPSI in this population, frequent, small set analyses were completed to monitor if this tool was psychometrically sound for use in this population. These early analyses provided evidence that although the EPSI had the potential for use in our target population, modifications were needed to enable children with motor delay to demonstrate their best performance (Dusing et al., 2019). The need to adapt the EPSI for use in children with motor delay is supported by previous outcome measures that were developed or adapted specifically for use in children with motor delay. Examples include the Gross Motor Function Measure that was developed to measure change in gross motor skills in children with CP and the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III) Low Motor/Vision Version which uses an adapted protocol to optimize the child's test performance in children with motor or visual impairments (Palisano et al., 1997; Visser et al., 2013).

In total, four major modifications over the course of five years were made to the EPSI, including to the assessment and behavioral coding protocols, problem-solving key skill definitions, and the use of a new weighted scoring model (Table 1). All modifications were grounded in theory, guided by empirical data, and reinforced by observational analysis or expert opinion. When taken together, the modifications resulted in a significant alteration to the EPSI. Thus, we refer to the combination of all these modifications as the Assessment of Problem-Solving in Play (APSP) for the remainder of this paper but acknowledge that the APSP is based on the EPSI (Table 1). The APSP has been shared with the Juniper Garden's team who developed the EPSI (J. Buzhardt, & D. Walker, personal communication, November 13, 2019).

Table 1. Comparison of two assessments of problem-solving.

	Early Problem-Solving Indicator (EPSI) (Greenwood et al., 2006)	Assessment of Problem-Solving in Play (APSP)
Subjects included in development and validation	30 children aged 12-48 months with "Typical Development"	124 children, 7-27 months of age with a Motor Delay (Harbourne et al., 2018)
Assessment	Traditional EPSI with no mention of subjects needing postural support	**Modification #1**(O'Grady & Dusing, 2016) Researcher provided adequate postural support to allow child to use arms freely without using a seat. Toys (popup, cups, gumball tower) selected based on least force required to solve.
Scoring	Scored live or from viewing videotaped assessment once	**Modification #2** (Dusing et al., 2015) Behavioral coded video taped assessment using Datavyu Coding Software
Problem-Solving Skills	4 skills: Look, Explore, Function, Solution (Greenwood et al., 2006)	**Modification #3** (Molinini et al., 2019) 5 skills: Look, *Simple Explore*, *Complex Explore*, Function, Solution
Weighted Scoring Model	Sequential ordering of skill scores: (Juniper Gardens Children's Project, 2017) • Look *1 • Explore *1 • Function *2 • Solution *3	**Modification #4** (non-published video observation & expert opinion) Larger weight assigned to more difficult skills: • Look *1 • Simple Explore *2 • Complex Explore *5 • Function *8 • Solution *16

Modification 1

Due to the activity limitations and young age of the participants included in our target population, many were unable to assume or maintain independent sitting to complete the assessment. A 22-month-old child who was unable to sit independently achieved different scores on the EPSI when tested in different positions (prone, supine, or supported sitting), with the highest scores and most advanced skills achieved in supported sitting (O'Grady & Dusing, 2016). This finding was in line with previous research that indicates children with typical or delayed motor development generally perform more complex and advanced manual exploration when in sitting as compared to supine or prone (Lobo et al., 2014; Soska & Adolph, 2014). Therefore, the assessment protocol was established to provide the least amount of postural support to allow the child to sit upright and freely use their hands during the assessment when using the EPSI in future research.

Modification 2

The first use of the EPSI to evaluate outcomes by our research team was on a cohort of infants born extremely preterm who were 3 months corrected age at their first EPSI assessment (Dusing et al., 2015; 2018). Since the EPSI had not been used on a sample that young, we needed to ensure that the problem-solving skill definitions as stated by the EPSI were applicable to this young population of infants. To do this we chose to score the videos using behavioral coding software (Datavyu 1.3.7), allowing more insight into the child's behavior which enabled the expansion of the EPSI skill definitions to fit this younger population. Expanded and more detailed definitions coupled with

behavioral coding permitted more precise identification of when each problem-solving skill started and stopped, allowing for duration of skills to be quantified in addition to frequency. As part of our target population were infants born preterm, it was important to capture both the frequency and duration of behaviors as previous research has indicated that this population tend to exhibit a high frequency of behaviors but at lower durations (Landry & Chapieski, 1988; Lobo et al., 2015; Sigman, 1976). Additionally, children born preterm and those with motor delay may spend more time not attending to the task and behavioral coding allowed us to capture this time as well, whereas the original EPSI did not have a score that captured when a child was not engaged with the toy (Landry & Chapieski, 1988; Rose et al., 2001). We found that the use of durations was insightful for young children born preterm, but when assessing slightly older children over longer periods of time, we found that behavioral coding only the frequency of problem-solving skills was sufficient to detect changes over time. However, the detailed nature of behavioral coding generally resulted in the identification of higher frequency counts which prevented comparison with the EPSI age norms (Juniper Gardens Children's Project, 2017).

Modification 3

Analysis of the 12-month trajectory of problem-solving in 8 children with typical development (5-12 months) and 24 children with motor delays (10-24 months) who were participating on the START-Play clinical trial, indicated that participants with similar motor skills (emergence of sitting) were spending the majority of the assessment in exploration (Marcinowski et al., 2019; Molinini et al., 2019). Therefore, the total problem-solving score was heavily dependent on the frequency of Explores. This finding was expected due to the developmental age of the participants and/or inherent limitations due to the motor delay. Surprisingly, a subset of the participants had their highest frequency of Explores at 6 months and the change from 6 to 12 months was either nonsignificant or declined without a complementary increase in Functions, even though their BSID-III cognitive scores were increasing (Dusing et al., 2019; Molinini et al., 2019). Visual inspection of the EPSI assessment videos identified that the participants were changing how they were exploring but this change was not being reflected in their EPSI score. The participants were shifting from modes of simple, high frequency exploration, such as mouthing or banging (Rochat, 1989), to more mature goal-directed, low frequency exploration including trying to fit pieces together or place small items inside larger items (Keen, 2011). However, these behaviors would both be scored as exploration in the EPSI and because they were not advancing to Functions, their score was stagnant. This led the research team to hypothesize that a modification to the Explore key skill may enhance the ability to detect subtle change in problem-solving skills in young children with motor delays.

To mitigate the impact of motor delay and acknowledge children's emerging understanding of the problem-solving steps needed to solve the toy, the EPSI problem-solving skill of Explore was divided into two mutually exclusive hierarchical skills; Simple Explore and Complex Explore (Figure 1). A Simple Explore maintained the same definition as Explore in the EPSI which is the child manipulates the toy to gain knowledge

Problem-Solving Skill	Definition	Example	Weighted Score
Look* 	Scored when the child gazes at the toy for greater than 3 seconds	Child looks at *pop-up, cups*, or *gumball* toy without touching it	1
Simple Explore 	Scored when the child manipulates the toy in a manner that provides information about the object properties	*Popup:* lifts the toy and shakes it *Cups:* mouths the cups *Gumball:* throws or bangs the balls together	2
Complex Explore 	Scored when the child attempts to execute a Function but is unsuccessful	*Popup:* pushes instead of twists button to pop up animal *Cups:* attempts to nest larger cup inside smaller cup *Gumball:* reaches inside tower to remove balls instead of pushing lever	5
Function* 	Scored when the child completes one step of a toy's function	*Popup:* pops up or pushes down one animal *Cups:* nests one cup inside another or stacks one cup on another *Gumball:* places one ball inside tower or pushes lever to remove one ball	8
Solution * 	Scored when the child completes all possible functions of the toy	*Popup:* pop's up and pushes down all 4 animals *Cups:* nests or stacks 4 cups in the correct order *Gumball:* releases 3 balls from the tower using the lever mechanism	16

Figure 1. Problem-solving skill definitions and examples. *=definitions come directly from the Early Problem-Solving Indicator (Greenwood et al., 2006) Notes: The popup toy in this analysis contained a push, slide, turn, and slide mechanism to pop up one animal. The cups used consisted of 5 round or square sequentially sized cups. The gumball toy consisted of a gumball tower and 5 balls, the tower contained a larger opening on the top for the child to place the balls, and a distal lever that the child pressed one time to release one ball from the tower.

about the object properties (Greenwood et al., 2006). A Complex Explore was defined by a child's unsuccessful attempt to execute a Function, thus demonstrating an emerging understanding of problem-solving. Following this modification, the data from the sample described in the previous paragraph (8 children with typical development and 24 children with motor delays) was re-analyzed to compare problem-solving scores using Explore only vs. using Simple and Complex Explore. Results indicated that the addition of Simple and Complex Explore may capture the subtle shift in children's

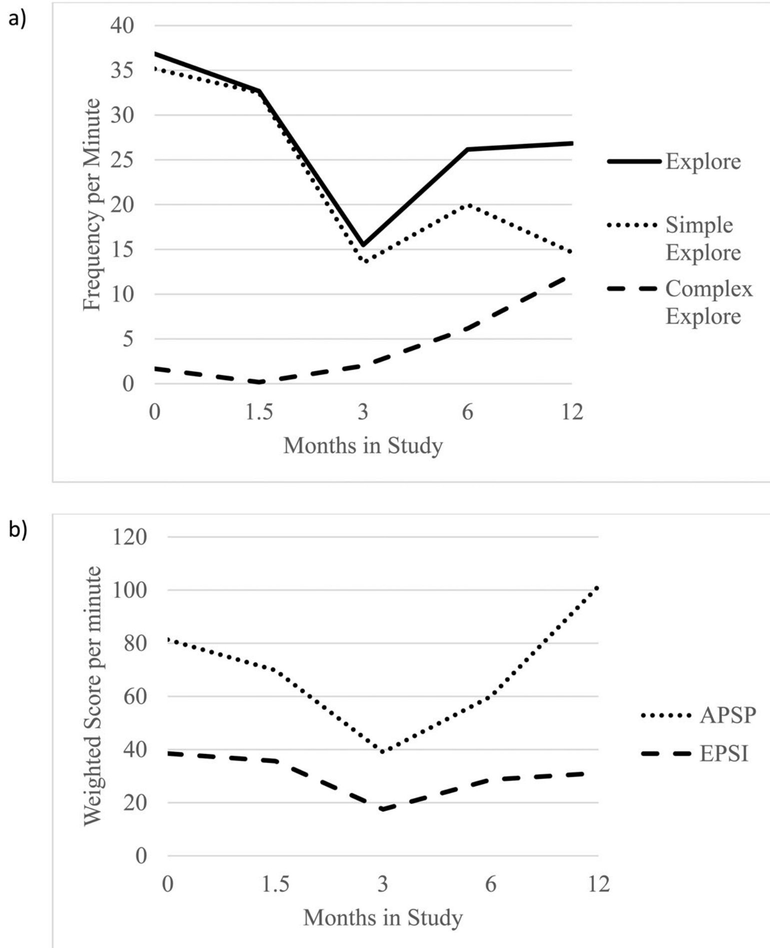

Figure 2. Comparison of the 12 month trajectory of Explore vs. Simple and Complex Explore (panel a) and EPSI vs. APSP weighted scoring models (panel b) in a child eight months old at baseline with a mild motor delay. Over 12 months, the frequency of Explores decreased (panel a) and no Functions or Solutions were performed (not shown). Subsequently, the EPSI weighted score decreased over time (panel b). After completing Modification's 3 and 4, the assessments were re-scored using Simple and Complex Explores and the newly developed APSP weighted scoring model. These modifications revealed an increase in Complex Explores and APSP weighted score over time.

emerging problem-solving skills. While Simple Explores did decrease over time in a subset of the participants, there was a complementary increase in Complex Explores (Figure 2, panel a) (Molinini et al., 2019).

Modification 4

The EPSI uses a weighted scoring model to create a single, total EPSI score that is derived from the frequencies of each problem-solving skill. The addition of Complex Explore required a new weighted scoring model. Consistent with the original EPSI weighted system, a Look was a weight of one as it is the least difficult skill (Table 1).

Our experience measuring looking, reaching, and manual exploration behaviors in young infants supported a score of two for Simple Explore. To establish the weights for the remaining skills, the research team reviewed a subset of assessment videos of participants with typical motor development or motor delay to identify if any trends were observable to guide this process. During the video review, it was noted that participants trialed on average, three types of Simple Exploration before advancing to Complex Explore, and trialed on average, three Complex Explore strategies before advancing to Functions. Therefore, we applied an interval of three between Simple and Complex Explore and between Complex Explore and Function. Lastly, Solutions were allotted a score that was double that of a Function, as it was our team's interpretation that a Solution was much harder than a Function. Based on these observations and our expert opinion, the weighted scoring model that included Complex Explore was: Looks$^*1+$Simple Explore$^*2+$Complex Explore$^*5+$Function$^*8+$Solution^{*16}. When comparing the performance of the new weighted scoring model to the EPSI scoring model (Table 1), it was clear that the new weighted scoring model, with the addition of Simple and Complex Explores, was able to provide more insight into identifying problem-solving skills that were slowly emerging (Figure 2, panel b).

After developing the APSP but prior to its use in research or clinical practice, the quality of the measurement properties needed to be assessed. Thus, the purpose of this study was to evaluate the construct validity and the responsiveness of the APSP in children with motor delay who were 7-27 months of age at assessment. The BSID-III was used as the comparison measurement tool for this study as it is the most widely used tool to assess cognition in young children with motor delay (Visser et al., 2013). We did not consider the BSID-III as the "gold standard" (and thus did not evaluate evidence of criterion validity), as it measures the broad construct of cognition and does not isolate the domain of problem-solving (Bayley, 2006).

Construct validity refers to the degree to which a test measures what it purports to measure (Mokkink et al., 2010). It was evaluated by estimating (a) correlations between the baseline APSP and BSID-III Cognitive Scale raw scores, as an indication of convergent validity evidence; and (b) mean differences in baseline APSP scores between known subgroups, defined by the infants' baseline severity of neuromotor delay, as an indication of discriminative or known-groups validity (Mokkink et al., 2010). We hypothesized that the APSP and BSID-III Cognitive Scale raw scores would have a strong correlation (>0.5) at baseline and that severity of motor delay would have a significant impact on APSP scores at baseline, with children with less severe motor delay having significantly more positive baseline scores.

In parallel, responsiveness indicates the longitudinal validity evidence of the APSP to detect important change over time (Mokkink et al., 2010). Responsiveness, based on the construct approach, was evaluated by estimating (a) correlations between the change in APSP and BSID-III cognitive scores over time; and (b) mean differences in the change in APSP scores over time between severity groups (Mokkink et al., 2010). We hypothesized that the change scores of the APSP and BSID-III cognitive raw scores over time would be strongly correlated (>0.5) and that all severity groups would demonstrate positive change on average but change scores would be significantly more positive for children with less severe motor delay.

Table 2. Participant characteristics.

	Mild Motor Delay (n = 64)	Moderate Motor Delay (N = 33)	Severe Motor Delay (N = 28)	Total Sample (N = 125)
Baseline Age (months)				
Mean	9.82	11.54	11.83	10.72
Min	6.93	7.08	7.21	6.93
Max	16.26	17.18	16.32	17.18
Sex				
Male	37 (57.8%)	19 (57.6%)	14 (50%)	70 (56%)
Female	27 (42.2%)	14 (42.4%)	14 (50%)	55 (44%)
Child race				
White	42 (65.6%)	22 (66.7%)	17 (60,7%)	81 (67.5%)
Black	5 (7.8%)	5 (15.2%)	5 (17.9%)	15 (12.5%)
Other	14 (21.9%)	4 (12.1%)	6 (21.4%)	24 (20%)
Gestational Age				
>=37 weeks	48 (75%)	19 (57.6%)	15 (53.6%)	82 (65.6%)
34-36 weeks	6 (9.4%)	1 (3.0%)	2 (7.1%)	9 (7.2%)
32-33 weeks	4 (6.3%)	4 (12.1%)	2 (7.1%)	10 (8%)
25-31 weeks	4 (6.3%)	3 (9.1%)	6 (21.4%)	13 (10.4%)
<25 weeks	2 (3.1%)	6 (18.2%)	3 (10.7%)	11 (8.8%)
Parent Report of History of brain injury				
Yes	8 (12.5%)	7 (21.2%)	16 (57.1%)	31 (26.3%)
No	53 (82.8%)	23 (69.7%)	11 (39.3)	87 (73.7%)
History of visual problems				
Yes	7 (10.9%)	9 (27.3%)	18 (64.3%)	34 (28.1%)
No	55 (85.9%)	22 (66.7%)	10 (35.7%)	87 (71.9%)
History of hearing problems				
Yes	8 (12.5%)	5 (15.2%)	6 (21.4%)	19 (15.7%)
No	54 (84.4 %)	26 (78.8%)	22 (78.6%)	102 (81.6%)
History of seizures				
Yes	9 (14.1)	10 (30.3%)	10 (35.7%)	29 (24.2%)
No	52 (81.3%)	21 (63.6%)	18 (64.3%)	91 (75.8%)
baseline BSID-III Cognitive scaled score				
Mean	7.2 (3.1)	4.9 (3.5)	1.9 (1.6)	5.4 (3.7)
min	1	1	1	1
max	13	14	7	14
Baseline Bsid-III Gross motor scaled score				
Mean	3.5 (2.1)	2.0 (1.8)	1.1 (.5)	2.5 (2.0)
min	1	1	1	1
max	8	7	3	8

Methods

Participants

Data were drawn from 125 infants, 7-16 months at baseline (Mean baseline age was 10.72, SD = 2.62 months) who were participating in a larger longitudinal randomized controlled trial that examined the efficacy of a novel physical therapy intervention (Table 2) (Harbourne et al., 2018). Children from both the intervention and usual care groups from the clinical trial were included in this analysis. Participants were enrolled if they scored greater than one standard deviation below the mean on the Bayley Scales of Infant and Toddler Development, Third Edition (BSID-III) Gross Motor Subtest at baseline, demonstrated an emerging ability to sit (defined as the ability to prop sit for 3 seconds, while unable to transition in and out of sitting independently), and had some active arm movements in sitting (Harbourne et al., 2018). Infants were excluded if they had a disability of a progressive nature or a medical complication limiting participation

in assessments (Harbourne et al., 2018). All participants qualified for and were not prohibited from receiving early intervention services during the study.

After baseline assessment, severity of motor delay (mild vs. moderate vs. severe) was assigned based on a scale incorporating a multitude of developmental characteristics used as part of the larger clinical trial in which the participants were enrolled (Harbourne et al., 2018). 51.2% of the sample identified as having mild motor delay (Mean baseline age 9.82, SD = 2.19 months), 26.4% moderate motor delay (Mean baseline age 11.54, SD = 2.89 months) and 22.4% severe motor delay (Mean baseline age 11.83, SD = 2.7 months) (Table 2). Per parent report, ten participants were diagnosed with CP, 1 Hemiplegia, 5 Tetraplegia and 4 Spastic CP. Twenty-four participants were later diagnosed with an unspecified motor delay and 12 were diagnosed with a global developmental delay of unspecified origin (Table 2).

Measures

BSID-III

All sections of the BSID-III were administered at each visit but only the Cognitive Scale raw score was used as an outcome in this study (Bayley, 2006). The BSID-III Cognitive Scale includes items on sensorimotor development, object relatedness, exploration and manipulation, concept formation, and memory (Bayley, 2006). The BSID-III Cognitive Scale has strong correlation with the Peabody Developmental Motor Scales (r = .91) (Gills et al., 2018). Three physical therapists who completed reliability training prior to beginning the study and were blinded to group assignment scored all BSID-III assessments from video tape. The BSID-III Gross Motor Subtest was completed at enrollment and average baseline scores are provided in Table 2, but these scores were not an outcome measure for this analysis. Twenty percent of all BSID-III Cognitive Scale assessments were rescored for inter- and intra-rater reliability demonstrating strong agreement within and between raters (inter-rater reliability: ICC = 0.994 (0.002 – 0.996); Intra-rater reliability: ICC = 0.999 (0.999-0.999).

APSP

APSP Assessment. The APSP is a 6-minute assessment of problem-solving, described previously, in which children interact with one set of 3 toys, each for 2 minutes. All assessments were completed in a sitting position without the use of a chair or highchair and the assessor provided the least amount of postural support the child needed to sit upright and freely use their hands. If the child was unable to visualize the toy on the ground due to poor head control, the assessor raised the surface of the play area on a small table or had a parent or second assessor hold the toy. The assessor acted as a play partner to the child and did not provide any insight as to how to interact with or solve the toy. Standardized verbal cues, such as "what can you make the toy do?" or "where do the balls go?" were used every 20 seconds to re-direct the child's attention to the toy if needed. In the rare case that the child refused any support from the assessor, the parent provided postural support with cues from the assessor.

The toys selected were validated as part of the EPSI development due to their ability to evoke Functions and Solutions (Greenwood et al., 2006) (Figure 1). Additionally, the

toys selected are common, inexpensive toys that a researcher would have access to and are easy to transport. In this study, care was taken to select versions of the toys used by the EPSI developers that required minimal postural control or force generation and could be explored with a single upper extremity. After two minutes, the assessor transitioned from one toy to the next using cues such as "bye-bye pop-up friends" and "wow, look at the gumballs". The efficient transition allowed for an average of seven minutes to complete the assessment, two minutes per toy with 30 seconds allotted for transition time between toys.

Scoring the APSP. To quantify problem-solving skills, coders marked the frequency in which 5 hierarchical problem-solving skills occurred using the open-sourced Datavyu behavioral coding software. Assessment videos were behaviorally coded in real time, half time, and reviewed frame by frame as needed to identify the moment each skill started and stopped. After marking the frequency, the duration of time the child spent performing each individual problem-solving skill was then recorded. The scoring of one assessment took on average 30 minutes. Duration has been used as part of previous analyses in a cohort of younger participants who were not included in this study, but duration results are not included in this analysis (Dusing et al., 2015). The frequency of each problem-solving skill was then entered into a weighted scoring model (Table 1). The sum of the weighted score was then converted to rate per minute to adjust for any variation in assessment length resulting from infant fatigue, behavioral state, or equipment failure during the data collection.

Five coders from Virginia Commonwealth University completed all APSP coding for this analysis over the course of four years. SCD (last author) was trained and checked for reliability by the developers of the Early Problem-Solving Indicator. Per the EPSI training protocol, reliable examiners can train others. Coders scored practice videos until they reached 90% intra- and inter-rater reliability. On average, coders required 4-6 weeks of training to achieve reliability standards. Three of the coders were trained Physical Therapists and two were research assistants with undergraduate training in psychology. Inter-rater reliability for all problem-solving skills ranged from ICC = 0.82-0.98. Intra-rater reliability was calculated using the equation [Agreed/(Agreed + Disagreed)] x100. The average intra-rater reliability was 96%.

Procedures

Infants were assessed with the BSID-III and APSP up to 4 times over 12 months, including baseline, and 3-, 6-, and 12-months post baseline (Harbourne et al., 2018). Assessment visits were usually completed within 1 home visit, lasting 1 to 1.5 hours and one caregiver was present during all procedures (Harbourne et al., 2018). Since this was a multisite clinical trial spanning 5 sites over 4 years, there were up to 2 assessors per site that completed all assessments. Assessors were blinded to group assignment and assessments were frequently reviewed for adherence to assessment protocol. All coders were blinded to group assignment and trained to 90% reliability [agreement = agree/(agree + disagree)*100] through a series of standardized practice videos. Twenty percent of the assessments were re-scored by the same or a different coder for intra- and inter-

Table 3. Descriptive statistics for the Assessment of Problem-Solving in Play (APSP) and Bayley Scales of Infant and Toddler Development-Third Edition (BSID-III), cognitive scale raw scores.

	Total			Mild Motor Delay			Moderate Motor Delay			Severe Motor Delay		
	N	M	SD	N	M	SD	N	M	SD	N	M	SD
APSP												
0mo	125	51.87	26.44	64	63.68	20.07	33	54.32	24.00	28	21.99	18.00
3mo	115	59.19	30.35	61	68.02	24.42	29	64.78	32.21	25	31.15	24.90
6mo	109	71.54	36.95	58	84.34	28.60	26	76.71	37.44	25	36.45	32.12
12mo	101	81.82	43.75	51	97.62	34.25	25	87.34	46.51	25	44.07	36.37
BSID-III Cognitive Scale Raw Score												
0mo	125	28.10	7.77	64	31.13	4.97	33	29.33	7.16	28	19.71	7.88
3mo	115	33.53	9.02	61	37.20	5.08	29	34.28	8.17	25	23.72	10.47
6mo	107	36.25	9.50	56	40.13	5.44	26	38.62	8.62	25	25.12	9.11
12mo	100	41.69	13.65	50	49.08	8.62	25	41.20	12.02	25	27.40	12.07

N = Number of Participants, M = Mean, SD = Standard Deviation; 0 mo = baseline and other assessments are labeled at months post baseline.

rater reliability. For additional clinical trial related procedures including intervention fidelity and description refer to the protocol or outcomes paper for the clinical trial (Harbourne et al., 2018; An et al., 2020).

Statistical Analysis

Multivariate linear mixed modeling (LMM) with APSP- and BSID-III -specific fixed effects of time and random residual, intercept, and time variances and covariances, was performed to estimate the associations between APSP and BSID-III Cognitive Scale raw scores. Univariate LMM, with fixed effects of time, severity, time by severity, and baseline adjusted age, and random residual, intercept, and time variances and covariances, was performed to estimate the trajectories of APSP and BSID-III Cognitive Scale raw scores and examine the mean differences in scores by severity group.

Data were analyzed in the SAS software environment using the MIXED procedure. This procedure capitalizes on all available data such that all infants were included in the analyses even if they had missing data at some time points. Estimation was carried out via restricted maximum likelihood, and denominator degrees of freedom were computed according to the Kenward-Rogers approximation. Statistical significance was set at $\alpha = .05$. Practical significance of the associations between APSP and BSID-III Cognitive Scale raw scores was assessed by the magnitude of the correlations, where Cohen (1988) suggests that $r = .10$, .30, and .50 represent small, moderate, and large effects, respectively. Practical significance of the change in scores across the 12 months of the study was indicated by standardized response means (SRMs). Finally, practical significance of the difference in APSP scores in each motor delay group was indicated by Hedges' g, where $g = .20$, .50, and .80 represent small, moderate, and large effects, respectively (Cohen, 1988).

Results

Table 3 provides a comparison of the average APSP and BSID-III Cognitive Scale raw score at each assessment time point, for the total sample, mild, moderate, and severe motor delay group.

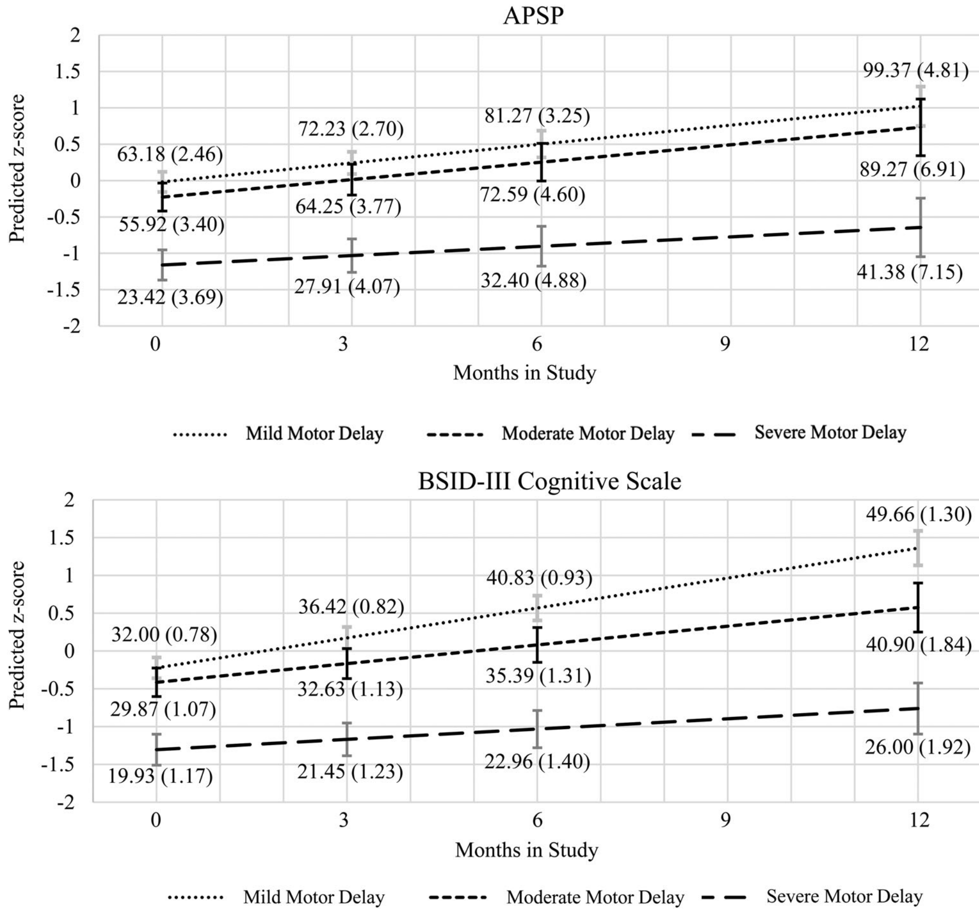

Figure 3. Model predicted trajectories of the Assessment of Problem-Solving in Play (APSP) and Bayley Scales of Infant and Toddler Development, Third Edition Cognitive Scale (BSID-III) z-scores and 95% confidence intervals. Predicted means and standard errors on the raw scores are provided next to the plotted points.

Construct Validity

Comparison with BSID-III

At baseline, BSID-III Cognitive Scale raw scores were significantly and strongly positively correlated with APSP scores ($r = .83$, $p < .001$).

Comparison between Severity Groups

At baseline, there was a significant omnibus effect of motor severity on APSP scores ($F[2, 123] = 39.74$, $p < .001$). There was no significant mean difference in APSP scores between participants with mild vs. moderate motor delay ($g = .34$, $p = .093$), but participants classified as having severe motor delay had significantly lower baseline scores on average than mild ($g = -2.02$, $p < .001$) and moderate ($g = -1.49$, $p < .001$) groups (Figure 3).

Responsiveness

Comparison with BSID-III

Linear change in BSID-III Cognitive Scale raw scores across time was significantly and strongly positively correlated with linear change in APSP scores across time ($r=.64$, $p < .001$).

Comparison between Severity Groups

Figure 3 shows the model-predicted trajectories of APSP and BSID-III Cognitive Scale raw scores across the 12 months of the study. Z-scores and corresponding 95% confidence intervals are plotted to facilitate comparison across measures, with predicted means and standard errors on the raw scales provided next to the plotted points. Significant and positive 0-12 month change was observed for all three severity levels for both APSP (mild: SRM $= 1.24$, $p < .001$; moderate: SRM $= 1.18$, $p < .001$; severe: SRM $= 0.75$, $p = .003$) and BSID-III Cognitive Scale raw scores (mild: SRM $= 2.13$, $p < .001$; moderate: SRM $= 1.61$, $p < .001$; severe: SRM $= 0.64$, $p = .003$).

There was a significant omnibus effect of severity of motor delay on linear change in APSP scores across time ($F[2, 104] = 3.40$, $p = .037$). There was no significant mean difference in linear change in APSP scores across time between participants classified as mild vs. moderate motor delay ($g = .07$, $p = .689$) or moderate vs. severe motor delay ($g = .36$, $p = .066$), but participants classified as severe motor delay exhibited significantly less change on average than participants with mild motor delay ($g = -.52$, $p = .012$) (Figure 3).

Discussion

The results of this study provide evidence for the construct validity and responsiveness of the APSP in quantifying problem-solving skills in young children with motor delay 7-27 months of age. The modifications made to the EPSI (Table 1) in response to previously published and unpublished data analyses proved beneficial as the newly developed APSP was able to capture change over time and discriminate between participants of different motor abilities.

Construct validity was measured two ways in this analysis; convergent and discriminative or known-groups validity. Convergent validity was evaluated by comparing baseline APSP and BSID-III cognitive raw scores at baseline. Results supported convergent validity in that the APSP scores were strongly correlated with the BSID-III ($r = .83$, $p < .001$) such that if a participant scored high on the APSP relative to other infants then s/he was also likely to score high on the BSID-III relative to other participants. The APSP scores also yielded discriminative validity evidence as the participants with severe motor delay scored significantly lower at baseline as compared to the participants with mild or moderate motor delay. This finding of differences in scores between the relevant subgroups at baseline confirms that degree of motor delay impacts problem-solving abilities and children with more severe motor delay will have, on average, lower problem-solving scores at individual timepoints (Figure 3).

Responsiveness was measured by comparing change of the APSP and BSID-III cognitive raw scores and the difference in change in APSP scores between motor severity subgroups over the course of 12 months. There was a strong correlation in APSP and BSID-III scores ($r = .64$, $p < .001$) such that if a participant demonstrated greater change on the APSP relative to other participants then s/he was also likely to demonstrate greater change on the BSID-III relative to other participants. The APSP was able to discriminate change scores between participants, highlighting the longitudinal impact of motor delay on trajectory of problem-solving skills. Children with more severe motor delay, on average, will demonstrate less change across time that children with more mild motor delay, indicating that motor skills impact both current and future problem-solving abilities (Figure 3).

Overall, the results support the use of the APSP as a research tool for young children with motor delay when comparing among different groups and evaluating associations between problem-solving and other constructs. Although the results of this paper establish important validity and responsiveness evidence, further research is needed to establish the clinical feasibility of the APSP and its use in other populations of children and those younger than 7 or older than 27 months.

Implications and Future Research

The administration of the APSP is efficient and practical for most therapy environments whether in a research lab, hospital, classroom, clinic, or home. The APSP can be completed with one assessor using a tripod for videotaping or with a second person videotaping. The assessment can be completed by a trained assessor or in the rare situation if the child is wary of the assessor or generally fussy, the APSP can be administered by a parent with cues from a trained assessor. The APSP uses 3 toys that evoke problem-solving behaviors and that are available at most clinics or readily available to inexpensively purchase.

However, there are time, training, and resource limitations in the APSP scoring scheme that hinder the generalizability of the measure for clinical use. The APSP scoring described in this paper uses Datavyu software and preprogrammed scripts that are open access and can be downloaded by any user to a Mac or PC. The Datavyu website is user friendly and provides tutorials allowing for coders with no previous coding experience to learn the coding system. However, some prior knowledge of Datavyu or other computer language is beneficial to expedite the learning process. Scoring the APSP relies on a computer with fast processing speed to complete detailed behavioral coding which may not be feasible in a clinical environment. At this time, although administration of the APSP is feasible in a clinical setting, scoring the measure is not. However, a study is underway that pilots a new scoring methodology on a sample of early interventionists or special education teachers with no previous training that may be more clinically applicable.

Limitations

While the results of this study support the use of the APSP, some limitations should be considered. One potential limitation of this analysis was the use of the BSID-III as the

comparative measure. Since the BSID-III measures cognition as a whole, it did not offer a direct comparison with the APSP. Though the BSID-III is widely considered the gold standard in measuring development in infants and toddlers, it is important to note that the cognitive assessment relies on completing motor tasks with timing demands which may result in many children with motor delays being "untestable" (Krumlinde-Sundholm et al., 2015; Morgan et al., 2019). The BSID-III has been found to have limitations in its ability to detect and track delays over time in children at risk for delays in development (Janssen et al., 2011; Lobo & Galloway, 2013; Morgan et al., 2019).

The repeated implementation of both the APSP and BSID-III could be viewed as a limitation of the study. The BSID-III is recommended for use no more than every 6 months to prevent learning of the tests, however this analysis uses data from a longitudinal study measuring change over time so the BSID-III was administered more frequently than recommended (Bayley, 2006; Harbourne et al., 2018). When the EPSI was developed, it was found to identify change during assessments performed every 3 weeks. As part of the longitudinal study in which data were drawn, the BSID-III and APSP were administered as frequently as 1.5 months apart. One potential consequence of repeated testing is the possibility that the participants learned the testing items. However, previous research found that repeated testing of infants with the BSID, First Edition did not bias the child's performance (Haskins et al., 1978). Additionally, the purpose of this paper was not to document the degree of delay, practice or learning effects, but rather to compare the abilities between two assessment tools repeated at the same frequency on the same group of participants. Any impact of test retention should be equal between each assessment and thus have the same impact.

Lastly, both the construct validity and responsiveness evidence are supported by the difference in scores between different severity groups. The allocation of participants to severity groups was based on objective data on the child's mobility and hand function from the baseline assessment (Harbourne et al., 2018). However, the severity classification may not be retained over time for all infants. This variability within and between groups is consistent with longitudinal development and we believe had limited impact on the interpretation of this data.

Conclusions

The results of this paper establish important evidence for the construct validity and responsiveness of the APSP for use in children with motor delays 7-27 months of age. Additionally, this paper highlights the importance of measuring the cognitive construct of problem-solving and provides a tool that can capture performance at a single visit as well as change over time.

Acknowledgments

We would like to acknowledge and thank Jay Buzhardt, PhD and Dale Walker, PhD from the Juniper Garden's Children's Project for their ongoing dialogue about measuring problem solving in children with motor disabilities and their review of an early draft of this paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Institute of Education Sciences, National Center for Special Education Research, Early Intervention and Early Learning in Special Education under Sitting Together And Reaching To Play (START-Play). (Award #R324A150103); Children's Hospital Foundation Research Grant under Relation Between Motor, Cognitive, and Language Skills during Infancy: An Extension of the START-Play Clinical Trial; Foundation for Physical Therapy under Supporting Play, Exploration, & Early Development (SPEEDI) for Infants Born Preterm: An Initial Efficacy Study.

Notes on contributors

Rebecca Molinini, PT is a Physical Therapist and PhD candidate in the Rehabilitation and Movement Science PhD Program at Virginia Commonwealth University.

Natalie A. Koziol, PhD is a research assistant professor in the Nebraska Center for Research on Children, Youth, Families and Schools at the University of Nebraska-Lincoln.

Tanya Tripathi, PhD is a post-doctoral researcher in the Pediatric and Rehabilitation Laboratory at The Ohio State University.

Regina T. Harbourne, PT, PhD is an Assistant Professor in the Rangos School of Health Sciences at Duquesne University. Her research interests include early postural control development, efficacy of early intervention, motor learning, and variability and complexity of motor control.

Sarah Westcott McCoy, PT, PhD is a Professor in the Department of Rehabilitation Medicine at the University of Washington. Her research interests include: 1) development of measurements for postural and motor control dysfunction; 2) evaluation of the effectiveness of specific motor learning interventions utilizing technology and gaming devices for children with developmental coordination disorder and cerebral palsy; 3) developing longitudinal data for impairments, health and participation and determining best service models for children with cerebral palsy.

Michele A. Lobo, PT, PhD is an Assistant Professor & Co-Director of the Move to Learn Innovation Lab in the Department of Physical Therapy and Biomechanics and Movement Science Program at the University of Delaware. Her research goal includes designing and testing assessments, interventions, and rehabilitation devices that maximize early learning and development.

James Bovaird, PhD is an Associate Professor of Educational Psychology at the University of Nebraska – Lincoln.

Stacey C Dusing, PT, PhD is a Professor in the Department of Physical Therapy at Virginia Commonwealth University. Her primary research interests include understanding the influence of motor control on global development and efficacy of physical therapy interventions for infants. Specifically, she focuses on the relationship between postural control, reaching, mobility, and cognition in infants with typical development and at risk for disabilities and interventions for infants born prematurely or with cerebral palsy.

ORCID

Rebecca M. Molinini http://orcid.org/0000-0002-8456-6460
Stacey C. Dusing http://orcid.org/0000-0003-3847-3893

References

Adolph, K. E., & Franchak, J. M. (2017). The development of motor behavior. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1-2), e1430. https://doi.org/10.1002/wcs.1430

An, M., Dusing, S. C., Harbourne, R. T., Sheridan, S. M., & START-Play Consortium. (2020). What really works in intervention? Using fidelity measures to support optimal outcomes. *Physical Therapy*, *100*(5), 757–765. https://doi.org/10.1093/ptj/pzaa006

Barnett, A. L., Guzzetta, A., Mercuri, E., Henderson, S. E., Haataja, L., Cowan, F., & Dubowitz, L. (2004). Can the Griffiths scales predict neuromotor and perceptual-motor impairment in term infants with neonatal encephalopathy? *Archives of Disease in Childhood*, *89*(7), 637–643. https://doi.org/10.1136/adc.2002.019349

Bayley, N. (2006). *Bayley scales of infant and toddler development, third edition: Administration manual*. Harcourt.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associations.

Dusing, S. C., Brown, S. E., Van Drew, C. M., Thacker, L. R., & Hendricks-Muñoz, K. D. (2015). Supporting play exploration and early development intervention from NICU to home: A feasibility study. *Pediatric Physical Therapy*, *27*(3), 267–274. https://doi.org/10.1097/PEP.0000000000000161

Dusing, S. C., Harbourne, R. T., Lobo, M. A., Westcott-McCoy, S., Bovaird, J. A., Kane, A. E., Syed, G., Marcinowski, E. C., Koziol, N. A., & Brown, S. E. (2019). A physical therapy intervention to advance cognitive and motor skills: A single subject study of a young child with cerebral palsy. *Pediatric Physical Therapy*, *31*(4), 347–352. https://doi.org/10.1097/PEP.0000000000000635

Dusing, S. C., Tanya Tripathi, T., Marcinowski, E. C., Thacker, L. R., Brown, L. F., & Hendricks-Muñoz, K. D. (2018). Supporting play exploration and early developmental intervention versus usual care to enhance development outcomes during the transition from the neonatal intensive care unit to home: A pilot randomized controlled trial. *BMC Pediatrics*, *18*(1), 1–12. https://doi.org/10.1186/s12887-018-1011-4

Gibbs, E. D., & Teti, D. M. (1990). *Interdisciplinary assessment of infants: A guide for early intervention professionals*. Brookes.

Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. *Annual Review of Psychology*, *39*(1), 1–42. https://doi.org/10.1146/annurev.ps.39.020188.000245

Gills, K., Osiovich, A., Synnes, A., Agnew, J. A., Grunau, R. E., Miller, S. P., & Zwicker, J. G. (2018). Concurrent validity of the Bayley-III and the Peabody Developmental Motor Scales-2 at 18 months. *Physical & Occupational Therapy in Pediatrics*, *39*(5), 514–524.

Greenwood, C. R., Carta, J. J., & McConnell, S. (2011). Advances in measurement for universal screening and individual progress monitoring of young children. *Journal of Early Intervention*, *33*(4), 254–267. https://doi.org/10.1177/1053815111428467

Greenwood, C. R., Walker, D., Carta, J. J., & Higgins, S. K. (2006). Developing a general outcome measure of growth in the cognitive abilities of children 1 to 4 years old: The early problem-solving indicator. *School Psychology Review*, *35*(4), 535–551. https://doi.org/10.1080/02796015.2006.12087960

Harbourne, R. T., & Berger, S. E. (2019). Embodied cognition in practice: Exploring effects of a motor-based problem-solving intervention. *Physical Therapy*, *99*(6), 786–796. https://doi.org/10.1093/ptj/pzz031

Harbourne, R. T., Dusing, S. C., Lobo, M. A., Westcott-McCoy, S., Bovaird, J., Sheridan, S., Galloway, J. C., Chang, H.-J., Hsu, L.-Y., Koziol, N., Marcinowski, E. C., & Babik, I. (2018). Sitting Together and Reaching to Play (START-Play): Protocol for a multisite randomized controlled efficacy trial on intervention for infants with neuromotor disorders. *Physical Therapy*, *98*(6), 494–502. https://doi.org/10.1093/ptj/pzy033

Haskins, R., Ramey, C. T., Stedman, D. J., Blacher-Dixon, J., & Pierce, J. E. (1978). Effects of repeated assessment on standardized test performance by infants. *American Journal of Mental Deficiency*, *83*(3), 233–239.

Janssen, A., Akkermans, R. P., Steiner, K., de Haes, O., Oostendorp, R., Kollee, L., & Nijhuis-van Der Sanden, M. (2011). Unstable longitudinal motor performance in preterm infants from 6 to 24 months on the Bayley Scales of Infant Development-Second edition. *Research in Developmental Disabilities*, *32*(5), 1902–1909. https://doi.org/10.1016/j.ridd.2011.03.026

Juniper Gardens Children's Project. (2017) *EPSI update*. Retrieved June 7, 2019, from https://igdi.ku.edu/epsi-updates/

Keen, R. (2011). The development of problem solving in young children: A critical cognitive skill. *Annual Review of Psychology*, *62*(1), 1–21. https://doi.org/10.1146/annurev.psych.031809.130730

Krumlinde-Sundholm, L., Ek, L., & Eliasson, A. C. (2015). What assessments evaluate use of hands in infants? A literature review. *Developmental Medicine & Child Neurology*, *57*, 37–41. https://doi.org/10.1111/dmcn.12684

Landry, S. H., & Chapieski, M. L. (1988). Visual attention during toy exploration in preterm infants: Effects of medical risk and maternal interactions. *Infant Behavior and Development*, *11*(2), 187–204. https://doi.org/10.1016/S0163-6383(88)80005-5

Lobo, M. A., & Galloway, J. C. (2013). Assessment and stability of early learning abilities in preterm and full-term infants across the first two years of life. *Research in Developmental Disabilities*, *34*(5), 1721–1730. https://doi.org/10.1016/j.ridd.2013.02.010

Lobo, M. A., Harbourne, R. T., Dusing, S. C., & McCoy, S. W. (2013). Grounding early intervention: Physical therapy cannot just be about motor skills anymore. *Physical Therapy*, *93*(1), 94–103. https://doi.org/10.2522/ptj.20120158

Lobo, M. A., Kokkoni, E., Carolina de Campos, A., & Galloway, J. C. (2014). Not just playing around: Infants' behaviors with objects reflect ability, constraints, and object properties. *Infant Behavior and Development*, *37*(3), 334–351. https://doi.org/10.1016/j.infbeh.2014.05.003

Lobo, M. A., Kokkoni, E., Cunha, A. B., & Galloway, J. C. (2015). Infants born preterm demonstrate impaired object exploration behaviors throughout infancy and toddlerhood. *Physical Therapy*, *95*(1), 51–64. https://doi.org/10.2522/ptj.20130584

Marcinowski, E. C., Tripathi, T., Hsu, L. Y., Westcott McCoy, S., & Dusing, S. C. (2019). Sitting skill and the emergence of arms-free sitting affects the frequency of object looking and exploration. *Developmental Psychobiology*, *61*(7), 1035–1047. https://doi.org/10.1002/dev.21854

Molinini, R. M., Inamdar, K., Marcinowski, E. C., Dusing, S. C., Koziol, N. K., Lobo, M. A., Harbourne, R., Westcott McCoy, S., & Bovaird, J. (2019). Measuring early problem-solving in children with motor impairments: Evaluating responsiveness over time using a modification of the early problem solving indicator. *Developmental Medicine & Child Neurology*, *61*(S3), 3–74.

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, *19*(4), 539–549. https://doi.org/10.1007/s11136-010-9606-8

Morgan, C., Honan, I., Allsop, A., Novak, I., & Badawi, N. (2019). Psychometric properties of assessments of cognition in infants with cerebral palsy or motor impairment: A systematic review. *Journal of Pediatric Psychology*, *44*(2), 238–252. https://doi.org/10.1093/jpepsy/jsy068

Newborg, J., Stock, J., Wnek, L., Guidubaldi, J., & Svinicki, J. (1984). *The Battelle Developmental Inventory: Examiner's manual*. DLM Teaching Resources.

O'Grady, M., & Dusing, S. C. (2015). Reliability and validity of play-based assessments of motor and cognitive skills for infants and young children: A systematic review. *Physical Therapy*, *95*(1), 25–38.

O'Grady, M., & Dusing, S. C. (2016). Assessment position affects problem-solving behaviors in a child with motor impairments. *Pediatric Physical Therapy*, *28*(2), 253–258.

Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E., & Galuppi, B. (1997). Development and reliability of a system to classify gross motor function in children with cerebral palsy. *Developmental Medicine and Child Neurology*, *39*(4), 214–223. https://doi.org/10.1111/j.1469-8749.1997.tb07414.x

Rochat, P. (1989). Object manipulation and exploration in 2- to 5- month- old infants. *Developmental Psychology*, *25*(6), 871–884. https://doi.org/10.1037/0012-1649.25.6.871

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2001). Attention and recognition memory in the 1st year of life: A longitudinal study of preterm and full-term infants. *Developmental Psychology*, *37*(1), 135–151. https://doi.org/10.1037/0012-1649.37.1.135

Sigman, M. (1976). Early development of preterm and full-term infants: Exploratory behavior in 8-month-olds. *Child Development*, *47*(3), 606–612. https://doi.org/10.2307/1128174

Soska, K. C., & Adolph, K. E. (2014). Postural position constrains multimodal object exploration in infants. *Infancy*, *19*(2), 138–161. https://doi.org/10.1111/infa.12039

Visser, L., Ruiter, S. A. J., Van Der Meulen, B. F., Ruijssenaars, W., & Timmerman, M. E. (2013). Validity and suitability of the Bayley-III Low Motor/Vision version: A comparative study among young children with and without motor and/or visual impairments. *Research in Developmental Disabilities*, *34*(11), 3736–3745. https://doi.org/10.1016/j.ridd.2013.07.027

Walker, D., Carta, J. J., Greenwood, C. R., & Buzhardt, J. F. (2008). The use of individual growth and developmental indicators for progress monitoring and intervention decision making in early education. *Exceptionality*, *16*(1), 33–47. https://doi.org/10.1080/09362830701796784

Walker, D., & Greenwood, C. (2010). The cognitive problem-solving IGDI, early problem-solving indicator. In J. Carta, D. Walker, C. Greenwood & D. Walker (Eds.), *Using IGDIs monitoring progress and improving intervention for infants and young children* (pp. 57–74). Paul H Brookes Publishing.

Yue, A., Jiang, Q., Wang, B., Abbey, C., Medina, A., Shi, Y., & Rozelle, S. (2019). Concurrent validity of the Ages and Stages Questionnaire and the Bayley Scales of Infant Development III in China. *PLoS One*, *14*(9), e0221675. https://doi.org/10.1371/journal.pone.0221675